# FORENSIC SPEECH AND AUDIO ANALYSIS FORENSIC LINGUISTICS

**1998 to 2001**

## A Review

**A.P.A. Broeders**

**Dept. of Handwriting, Speech and Document Examination**
**Netherlands Forensic Institute**
**Ministry of Justice**
**P.O. Box 3110**
**2280 GC RIJSWIJK**
**The Netherlands**

## ABSTRACT

**Although the development of state-of-the-art speaker recognition systems has shown considerable progress in the last decade, performance levels of these systems do not as yet seem to warrant large-scale introduction in anything other than relatively low-risk applications. Conditions typical of the forensic context such as differences in recording equipment and transmission channels, the presence of background noise and of variation due to differences in communicative context continue to pose a major challenge. Consequently, the impact of automatic speaker recognition technology on the forensic scene has been relatively modest and forensic speaker identification practice remains heavily dominated by the use of a wide variety of largely subjective procedures. While recent developments in the interpretation of the evidential value of forensic evidence clearly favour methods that make it possible for results to be expressed in terms of a likelihood ratio, unlike automatic procedures, traditional methods in the field of speaker identification do not generally meet this requirement. However, conclusions in the form of a binary yes/no-decision or a qualified statement of the probability of the hypothesis rather than the evidence are increasingly criticised for being logically flawed. Against this background, the need to put alternative validation procedures in place is becoming more widely accepted.**

**Although speaker identification by earwitnesses differs in some important respects from the much more widely studied field of eyewitness identification, there are sufficient parallels between the two for speaker identification by earwitnesses to benefit greatly from a close study of the guidelines that have been proposed for the administration of line-ups in the visual domain. Some of the central notions are briefly discussed.**

**Rapid technical developments in the world of telecommunications in which speech and data are increasingly transmitted through the same communication channels may soon blunt the efficacy of traditional telephone interception as an investigative and evidential tool. The gradual shift from analogue to digital recording media and the increasingly widespread availability of digital sound processing equipment as well as its ease of operation make certain types of manipulation of audio recordings comparatively easy to perform. If done competently, such manipulation may leave no traces and may therefore well be impossible to detect.**

**Authorship attribution is another forensic area that has had a relatively chequered history. The rapid increase in the use of electronic writing media including e-mail, sms, and the use of ink jet printers at the expense of typewritten and to a lesser extent hand-written texts reduces the opportunities of authorship attribution by means of traditional document examination techniques and may create a greater demand for linguistic expertise in this area.**

**A survey is provided of ongoing work in the area, based on reactions to a questionnaire sent out earlier this year.**

## INTRODUCTION

**The field of forensic speech and audio analysis comprises a wide range of activities of which the most spectacular is no doubt speaker identification. Other activities in the field include intelligibility enhancement of recorded speech samples, the analysis of disputed utterances, and the examination of the authenticity of audio recordings. A related though in many ways very different activity is linguistic authorship identification, the linguistic analysis of a spoken or written text undertaken with a view to establishing the identity of the author of that text.**

## SPEAKER IDENTIFICATION

**In spite of the regular appearance of high-tech speaker identification equipment in contempo-rary fiction and the film industry - as witness Tom Clancy's *Clear and Present Danger*, TV classics like *Star Trek*, *Charlie's Angels* and *Night Rider,* and perhaps to a lesser extent Alexander Solzhenitsyn's novel *The First Circle* - forensic speaker identification at the beginning of the 21st century remains an extremely challenging field, in which the promise held by technological advance is still largely unfulfilled. This applies even more strongly to large-scale forensic applications, which are as yet virtually non-existent. Outside the forensic arena, the introduction of automatic speaker identification technology in real-world applications has also fallen far short of what might be expected on the basis of its popular appeal and the promise the technology initially seemed to hold. So far it has mainly been limited to relatively low-risk applications, frequently involving communication by telephone. One of the more successful applications appears to be home incarceration surveillance [1]. Interestingly, the home detainee has no alternative to using the technique other than becoming an inmate again and will therefore normally tend to adopt a co-operative attitude, thereby fulfilling a major prerequisite for a successful operation of the technique. The recent implementation of a free speech speaker verification system by the First Direct Bank of Israel makes this subsidiary of Leumi Bank Group the first financial institution known to have introduced speaker identification technology for external clients [2]. Whether this will also turn the bank into a commercial success is a question that will be of more than passing interest to the speaker identification community. The supplier of the system, Comverse Technology, Inc. in Israel, has also developed a speaker identification facility as an add-on to a telephone interception system. Although the performance figures for the system specified in the documentation are quite high, the available specifications do not permit a meaningful assessment of its performance under real-world conditions. Nor is it clear whether the system is in fact in operational use.**

**Technically, a distinction is often made between *speaker recognition* - which is used as a cover term for the wide variety of situations in which people are identified, or strictly speaking individualised, on the basis of the sound of their voices - and the terms *speaker identification* and *speaker verification*. Identification systems are those which compare a test speaker against all the voices in a particular database to determine his or her identity; verification systems compare the test sample with a reference sample of the speaker who is claimed to have produced the test sample. In this sense, the forensic application typically amounts to a verification task, in that the question that needs to be answered tends to be whether the recorded voice is that of a particular speaker (i.e., the suspect). Occasionally, the term *authentication* is used for the process of establishing the identity of a speaker, but this term is generally more appropriately used to describe the examination of audio recordings with a view to establishing their authenticity (see 3.2 below).**

**The first and oldest form of speaker identification is of course speaker identification by ear- witnesses. The second major category comprises all forms of speaker identification by experts. At present, experts working in the field of forensic speaker identification use one of three approaches: (i) a phonetic-acoustic approach, (ii) a (semi-)automatic, analytical acoustic approach frequently combined with an auditory phonetic analysis, and (iii) a global automatic approach. Also methods are employed in which elements of the three types are combined in various ways.**

***SPEAKER IDENTIFICATION BY EARWITNESSES***
***Some history***

**The earliest recorded example here probably goes back to the Bible, where the book of *Genesis* relates how Jacob obtained the right of primogeniture from his elder brother Esau in return for a plate of lintel soup. Even though Isaac correctly recognised his younger son Jacob by his voice when he fraudulently presented himself to his father in the guise of his elder brother Esau, Isaac apparently allowed his sense of hearing to be overruled by his sense of touch and eventually gave Jacob his blessing. As such, the episode not only serves to illustrate that voice identification can be a powerful identification tool but it also provides a forensically relevant illustration of the notion that an increase in information does not necessarily lead to more knowledge.**

**While speaker identification evidence has been accepted by English courts since at least 1660 [3], one of the first highly publicised and most memorable applications of earwitness identification probably occurred as part of the Lindbergh baby kidnapping case in the 1930s. Almost three years after the event, the famous aviation pioneer claimed he recognised the German-accented, English speaking voice of the suspect as that of the abductor of his child. Misgivings about the validity of the identification by Lindbergh gave rise to the first systematic study of speaker identification by humans [4], which, though limited in scope and design [5], nevertheless produced interesting findings which partly inspired later research.**

*The present*

**Today, procedures in speaker identification by witnesses for evidential purposes typically involve the use of line-ups, following existing practice in the related domain of visual identification of persons by witnesses [6, 7, 8, 9]. It is worth stressing that the use of single person identification procedures, while producing positive identification scores in controlled experiments that are comparable to those obtained for (multi-person) line-ups, is generally rejected except to confirm an earlier identification. The reason is that line-ups, unlike identification procedures involving a single speaker only (i.e., the suspect), make it possible to detect the vast majority of false identifications. Detection is possible because in a properly designed line-up all members should have an equal probability of false identification, which reduces the risk of a false identification going undetected to 1/N in an N-person line-up, where 1/N is the likelihood of a false identification involving an innocent suspect. By contrast, there is of course no way in which false-positive identifications can be distinguished from correct identifications if only a single speaker is presented to the witness: both correct and incorrect identifications amount to selection of the suspect.**

**Compared with the rich literature on eyewitness identification, which has created sufficient consensus in the scientific community for a set of common guidelines to be formulated [10, 11], empirical studies of earwitness identification are few and far between, with the notable exception of the work of Yarmey [12, 13]. The last decades have seen various attempts to formulate guidelines for speaker identification by witnesses [6, 7, 14]. As in the visual domain, the main purpose of these guidelines is to control variables that might unduly affect the result of an identification test, rendering its meaning essentially null and void. In order to prevent such undesirable effects from occurring, procedures should be carefully thought out and strictly enforced. It is only when a positive identification cannot be argued to be due to other factors than an observed correspondence between the memory trace of the perpetrator's voice in the witness's memory and the sound of the suspect's voice that it can meaningfully contribute to the resolution of an identity question.**

**In addition to the 'formalised type' of speaker identification by witnesses that is used as evidence for or against an individual's involvement in a crime, there is of course a very much greater volume of speaker identification work carried out by members of police forces and interpreters involved in the processing of the vast quantities of telephone interceptions undertaken within the legal framework of various countries. However, speaker attributions in these calls are rarely made on the basis of the speaker's voice quality and speech patterns alone. Frequently, information about the line or number being intercepted and prior knowledge about the whereabouts of callers as well as information from earlier calls will play a major role in attributing a certain call to a certain speaker. In view of the vast amounts of telephone calls that are intercepted in some countries it is remarkable that relatively few challenges of these attributions are made and that an even smaller number of these challenges is successful.**

**A brief and concise review of the present state of the field of voice identification by witnesses is provided by Bull & Clifford [15]. There remains a great deal of research to be done to increase our insight in the effect of so-called estimator variables on speaker identification performance by earwitnesses. These estimator variables include the nature of the speech and of the voice quality of the speaker, the amount of speech heard by the witness, the delay between exposure to the voice of the unknown speaker and that of the suspect, the effect of telephone quality speech, of differences in age, gender and ethnicity between speaker and witness and of differences in communicative context. But also in the area of the so-called system variables, which are essentially under the experimenter's or forensic examiner's control, there is a need for considerable empirical research. A more fundamental problem is that of the questionable relevance of laboratory experiments to actual forensic casework. Casual observers in a non-threatening environment may well behave very differently from witnesses or suspects who are paying close attention to the person they are confronted with. However, it is clear that ethical considerations frequently stand in the way of attempts to accurately recreate real-world situations in a controlled experiment. Meanwhile, we would do well to heed the repeated warnings by Bull& Clifford [15], Yarmey [12] and others to treat speaker identification based on earwitness evidence with considerable caution.**

***SPEAKER IDENTIFICATION BY EXPERTS***
***Some history***

**The second and probably most frequently practised form in the forensic context is speaker identification by experts. By far the best-known pronunciation expert in the world of fiction is no doubt Professor Henry Higgins of *Pygmalion* and *My Fair Lady* fame, a character created by the Irish-English author G.B. Shaw. Possibly partly motivated by his experience as an Irish-accented speaker of English living in England, Shaw took a keen interest in matters relating to accent and dialect variation. It has often been suggested that he may have derived the inspiration for the character of Henry Higgins from professor Henry Sweet, a professor of linguistics in the University of London, whose ear was reported to be so acute that it allowed him to locate any Londoner within a radius of two or three miles of his home on the basis of his accent. Recently however, a rival model for Higgins has been claimed in the person of the even more renowned Daniel Jones, one of the pioneers of phonetic science and holder of the first chair of Phonetics in Britain [16]. Forensic applications of this type of speaker identification date from the first half of the last century, when the tape recorder and the sound spectrograph first made it possible to capture, replay, visually represent and analyse the inherently transient phenomenon of human speech.**

**One of the early approaches based on the use of the spectrograph initially showed considerable promise and came to be known as the voiceprint technique. This method was actually developed during the Second World War and essentially amounts to a visual comparison of spectrograms of linguistically identical utterances to determine whether they originate from a single speaker. In the second half of the last century the limitations of this approach were demonstrated to be so severe that its status soon became extremely controversial. While highly suggestive, the parallel with the fingerprint that the term voiceprint invokes, was shown to be utterly misleading. Unlike, fingerprints, or friction ridge patterns, spectrographic representations of speech are not invariant over time but highly variable within speakers, reflecting the inherent within-speaker variability that is characteristic of speech. Yet, in spite of the publication of a critical review of the use of the sound spectrograph for the purposes of forensic speaker identification carried out by a National Research Council committee of the American National Academy of Sciences [17], testimony based on modified forms of the voiceprint technique as practised by members of the VIAAS (Voice Identification and Acoustic Analysis Subcommittee) of the IAI (International Association for Identification) and others continues to be admitted as evidence in US courts of law. Saks [18] reports that by his last count it is admissible in 6 states, excluded in 8, admissible in 4 federal courts and excluded in 1. Brief surveys of the history of forensic speaker identification can be found in Braun & Künzel [19] and Meuwly [20].**

## *THE PRESENT*

**There are probably few forensic disciplines that are characterised by such a diversity of methods and procedures as the field of forensic speaker identification by experts. Basically, practitioners can be divided in three groups. The first group consists of trained phoneticians. They rely primarily on a combination of auditory phonetic analysis and a variety of acoustic measurements, and will generally only consider themselves competent to analyse speech samples in their own native language. Experts working within this phonetic-acoustic tradition, which was pioneered by the German BKA, are found in several government forensic laboratories including laboratories in Germany, Austria, Sweden, the Netherlands and Spain, and in private practice in countries like the United Kingdom and Germany. Perhaps the main criticism of this type of approach is that it has a strong subjective element and does not easily lend itself to validation [19].**

**The second group consists of those who use a set of semi-automatic measurements of particular acoustic speech parameters such as vowel formants, articulation rate and the like, sometimes combined with the results of a detailed, largely auditory phonetic analysis by a human expert. Examples of this type of approach are the methods used in Italy (RCIS), the *Dialect* system used in Russia (FSC) and Belarus, the *SIVE* system used in Lithuania, and the type of method that is frequently referred to as phonoscopy and is used in several Eastern-European countries, where it.**

**The third, most recent approach differs form the first two in that it is both automatic and global. It is automatic in the sense that any subjective analysis or evaluation of the speech material is reduced to a minimum; it is global in the sense that it does not address specific acoustic speech parameters but treats the signal as a physical phenomenon, more specifically as a continuously varying complex vibration. Most automatic speaker identification systems today use a form of Gaussian mixture modelling to characterise or 'model' the speech of the known, target speaker (i.e., frequently the suspect in a forensic application) and that of the unknown speaker (i.e., the perpetrator). In addition to this, a relevant speaker population is defined and a probability-density function of the speech variance of this set is calculated. What the method essentially sets out to do is determine how likely a degree of similarity or difference as found between the target speaker (say the suspect) and an unknown speaker (say the perpetrator) is to occur within the relevant population.**

**There are two main problems with this approach. One is general and applies equally to other types of speaker identification, the other is specific to a forensic-type application. The first is the problem of within- and between-speaker variation. In the context of automatic speaker verification this means that speaker models may overlap because they may occupy similar spaces in the mode of representation utilised by the automatic technique. As a result, speakers may not always be reliably distinguished, and the system will produce a certain proportion of false-positives. As this is precisely the type of error that the criminal justice system should always take great pains to avoid, the solution would seem to lie in adopting a more conservative decision criterion. However, as in all biometric identification techniques, there is a trade-off between false-positives and false rejections, which means that a system that is biased towards reducing false-positives will tend to produce unacceptable levels of false rejections and/or report unrealistically low probability scores for matches.**

**The second problem is related to the extreme sensitivity to transmission channel effects of automatic procedures, including the effects of different handsets, telephone lines, GSM-coding and perception-based compression techniques as used in Minidisk players and compression formats like MPEG. Recent research by Schmidt Nielsen & Crystal [21] confirms that, while human listeners show tremendous individual variability in performance, on average they tend to slightly outperform current state-of-the-art speaker verification systems. More importantly, they found that it is especially when conditions deteriorate as a result of differences in transmission channels, the presence of background noise and the like that human listeners are clearly superior to automatic speaker verification algorithms. It is precisely these conditions that tend to prevail in the forensic context.**

**Many observers of the scene believe that in order for the performance of automatic speaker recognition techniques to improve significantly, a better understanding of the speaker-specific, linguistic element in the speech signal will be necessary [1, 22, 23]. In recent years, what progress has been made, has been the result of an increasingly more effective exploitation of the information contained in the parameters extracted from the speech signal. The more fundamental type of research into what parameters truly capture speaker-specific information in the signal has received comparatively little attention, partly because this type of research is largely funded by application-oriented organisations and industries.**

**Fully automatic systems are gradually being introduced in forensic casework albeit on a relatively small scale. They are currently used in France [24] and Switzerland [20, 25], and are being tested in Spain [26] and the United States of America [27]. The FBI recently completed an evaluation project in which four automatic speaker recognition systems - out of a total of twelve systems whose developers were originally approached - were tested on a specially designed forensic database compiled by the FBI. The results confirm findings reported elsewhere in the literature that, whilst performance levels of automatic systems can be quite high when text and transmission conditions are controlled, deterioration tends to be dramatic in conditions resembling those usually encountered in the forensic domain. In an attempt to address the needs imposed by the forensic context, the FBI has designed a PC-based forensic automatic speaker recogniser (FASR), which outputs a log likelihood ratio score and a True/False decision. It also provides a measure of confidence for each recognition decision based on statistics with known error rates generated from large sample populations. There is no indication that the system is likely to be used for evidential rather than investigative purposes in the near future.**

*EXPRESSING CONCLUSIONS IN SPEAKER IDENTIFICATION*

**As in many other forensic identification disciplines, the formulation of the conclusions has been receiving considerable attention in the literature in the last few years [28, 29]. Traditionally, forensic speech experts, like their colleagues in other forensic disciplines, have been expressing their conclusions in terms of the probability that the questioned (trace) material originated from a given source, usually the suspect. In recent years, partly accelerated by the advent of DNA evidence, this type of conclusion has been challenged as logically flawed [30, 31]. Rather than reporting the probability of the questioned (speech) material originating from the suspect, the expert should report the probability of the evidence under two rival assumptions. One is the prosecution hypothesis: the assumption that the material originates from the suspect. The other hypothesis will generally be that the trace material originates from some other member of a potential suspect population, like the adult male population of a town or a particular region. The ratio between these two probabilities is called the 'likelihood ratio' and takes the form of a number, which in the case of DNA evidence, where known distribution frequencies of what are taken to be independent characteristics are multiplied, frequently assumes astronomical proportions. Yet, even these high numbers do not indicate how likely the trace material is to have originated from the suspect. This question may only be answered by the decision-making judge or jury, who are in possession of all known facts of the case, and is an - ultimate - issue that is considered outside the province of the expert.**

**What the likelihood ratio does express is the relative strength of the evidence, i.e., the extent to which it serves to make the prosecution case stronger or weaker. Because of the conceptual problems posed by the (large) numbers involved in the expression of the likelihood ratio, advocates of this approach, frequently known as Bayesians, have suggested the use of verbal scales [32]. It is worth stressing that these verbal terms express the relative strength of the evidence in favour of one proposition versus another and do *not* address the probability of the issue. Interestingly, methods involving automatic speaker identification algorithms such as those pioneered by Marescal [24], González-Rodriguez [26], Meuwly & Drygajlo [33] and Boves & Koolwaaij [34] easily lend themselves to the Bayesian approach and typically employ the likelihood ratio format to express their conclusions.**

**It has been argued that, in spite of the semblance of qualified opinion that the phrasing of their conclusions conveys, experts using traditional probability scales are effectively giving categorical judgements [35], without necessarily always being aware of this. What is clear is that they are generally unable to quantify their findings to the extent that the calculation of likelihood ratios becomes a realistic scenario. As in many other forensic identification disciplines, validation of the methods and of the resulting opinions is often lacking [18] and generally difficult to undertake. An interesting approach is that proposed by Found & Rogers [36] for forensic handwriting experts, which essentially treats the forensic expert as a black box whose performance can be measured by means of a specially designed comprehensive testing set. What is particularly intriguing is that a first analysis of their findings suggests that factors like the handwriting examiner's experience, training or age do not correlate with performance. On the basis of the data released so far though, it would appear that the failure to find a correlation may also be due to a ceiling effect in the test material.**

*ORGANISATIONS*

**In addition to VIAAS (the Voice Identification and Acoustic Analysis Subcommittee of the IAI), whose membership is predominantly American and is open only to those who are certified IAI members, there are currently two more international organisations whose members are in one way or another involved in forensic speaker identification. One is IAFP, the International Association for Forensic Phonetics (www.iafp.net), which was formally established in 1991 with the aim of providing a forum for those working in the field of forensic phonetics as well as ensuring professional standards and good practice in this area. Its membership is predominantly European and open only to established phoneticians. The second is the Expert Working Group for Forensic Speech and Audio Analysis. It had its inaugural meeting in Voorburg, the Netherlands in July 1998 and has since met in Madrid, Cracow and Paris, earlier this year. It forms part of ENFSI, the European Network of Forensic Science Institutes, which was set up in 1991, currently has 46 member laboratories in 31 countries and has been the driving force behind the establishment of Expert Working Groups for the various forensic disciplines. The Forensic Speech and Audio Analysis Working Group's membership includes experts from 17 European countries, as well as Turkey. One of the first priorities the Working Group has set itself is to collect information about the various procedures that are used in the member laboratories**

**In Spain, interest in forensic speaker identification and forensic acoustics is such that a national society, SEAF (Sociedad Española de Acústica Forense), was formally established in 2000. It brings together leading experts from such diverse fields as linguistics, electrical engineering, acoustics and forensic phonetics, with a view to improving methods and techniques in speaker identification and acoustic phonetics. In response to a recent, highly publicised court case in France, the Groupe Francophone de la Communication Parlée (GFCP) of the Société Française d'Acoustique (SFA), a group of predominantly French acousticians who are active in the field of speech technology, is currently circulating a petition on the internet demanding that voice expertise is no longer used by the legal system until such time as it is scientifically validated [37, 38]. It has earlier gone on record as arguing that it is unethical for anyone to be active in the field of forensic speaker identification without first demonstrating his or her competence in the field [39]. Both Braun & Künzel [19] and earlier Broeders [40] have argued that while there is a real concern that voice evidence is presented in an irresponsible and incompetent manner, the charge of unethical conduct is unfounded and the call for speech experts to dissociate themselves from forensic examinations will ultimately only result in an increased danger that phonetically uninformed testimony will go unchallenged. What is of the essence of course is that those who are involved in deciding issues of guilt, i.e., judges and juries, are made aware of the limitations of the methodology employed.**

AUDIO ANALYSIS
*INTELLIGIBILITY ENHANCEMENT*

**Although the use of dedicated filtering hard- and software is widespread in the latter type of work, the net effect of the use of this equipment in terms of getting additional words down on paper is not always impressive. In fact, a large proportion of the work carried out under this heading is probably primarily of a cosmetic nature; in judiciaries with a jury system in particular it is often necessary for all relevant speech recordings to be played in court. Removing unpleasant noises may facilitate listening for uninitiated listeners like members of the jury; it may also reduce fatigue and thereby increase productivity in those who have to transcribe large quantities of speech recorded under forensic real-world conditions.**

**The enhancement of clandestine or covert recordings, other than those made by private citizens, is not a core activity for many forensic laboratories for the simple reason that covert recordings made by police or other investigative forces will not normally be ruled admissible by a criminal court of law. Information obtained from such recordings cannot therefore be used for evidential purposes. The extent to which information obtained from enhanced audio recordings may play a role as an investigative tool and the efficacy of covert recording is hard to assess because by definition these matters do not lend themselves to public scrutiny. The public image of this type of activity is strongly shaped by publications like *Spycatcher* [41] and the Francis Ford Coppola film *The Conversation* (1974), in which Gene Hackman plays an audio surveillance expert who is slowly caving in under the psychological pressure of his job.**

**To achieve the best results in transcribing questioned utterances in low to extremely low quality recordings the use of highly competent and educated native speakers of the language variety in question is strongly recommended. A thorough familiarity with the accent and dialect of the speakers in the recording, as well as some familiarity with the details of the case, will often enable the analyst to compensate for the loss of redundancy of linguistic cues that is characteristic of poor quality recordings.**

***INTEGRITY AND AUTHENTICITY EXAMINATIONS OF AUDIO RECORDINGS***

**An interesting development in the field of authenticity and integrity examinations of audio recordings in the analogue domain is the use of Faraday crystals as pioneered by a number of Russian scientists. This may well turn out to be a welcome complement to the existing array of techniques in this field [42]. Traditionally, these include visual inspection of the tape and its housing, auditory analysis of the recording, magnetic development of the magnetisation patterns on the tape track, narrow band spectrum analysis of the recorded signal, and, last but not least, high resolution waveform analysis of the signal [43]. The analysis of replay transients, as Dean [44] calls them, plays a central role in these examinations. They may frequently shed light on the way in which the recordings on a questioned tape were made and may help establish the order in which these recordings were made. Unfortunately, there is still a relative dearth of experimental data on the reliability, robustness and consistency of replay transients of different tape recorders and there is still considerable uncertainty about the extent to which they may be used to identify individual analogue audio recorders. However, the new visualisation techniques may well reveal characteristics with the degree of detail that is required to improve the discriminatory power to the extent where it may be possible to trace a particular recording to a particular source recorder rather than merely to a particular brand and type.**

**In spite of this new development, overall prospects for this particular branch of forensic audio analysis are not too bright. The increasingly widespread availability of relatively inexpensive digital sound processing equipment and its ease of operation make certain types of manipulation comparatively easy to perform. If done competently, such manipulation may leave no traces and might therefore be impossible to detect from an engineering point of view. Failure to find positive evidence of copying and/or manipulation does not therefore imply that the recording under investigation must be a complete and uninterrupted magnetic registration of the acoustic events it is supposed to represent. Faced with recordings of extremely incriminating telephone conversations which were only available as copies, defence experts have been known to turn this argument round: if the recording is a copy it cannot be authenticated and must therefore be viewed with a high degree of suspicion regarding its authenticity. Not unnaturally, defence lawyers will pick this up and argue that this means that any recording that is not claimed to be an original recording but a copy should be ruled inadmissible as evidence because there is no way in which its integrity can be established. However, the mere fact that a recording is a copy does not ipso facto make it likely to have been tampered with.**

**The reservations made with respect to the authentication of analogue audio recordings apply even more strongly to digital audio recordings. These are becoming more numerous as digital dictation machines are becoming more and more common. As part of the chain-of-custody process, audio recordings, like all digital data, are increasingly required to be authenticated by means of checksums, hash codes or other methods to ensure their integrity.**

### *DISPUTED UTTERANCES*

**There are relatively few reports of work undertaken in this area. French [45] provides an illustration of some of the procedures that may be helpful here.**

**A related issue is the growing demand for speech recognition systems to meet the need to transcribe enormous quantities of forensic speech recordings. At present, the vast quantities of recorded speech generated by telephone interception systems are transcribed by relatively highly paid and trained human listeners. Most commercially available speech-to-text systems require extensive learning sessions, a (single) co-operative speaker and relatively high quality recordings to meet acceptable performance standards and are therefore unsuitable for forensic use. Interestingly, the Lithuanian Institute of Forensic Examination in Vilnius reports a system called Transcriber, produced by the Speech Technologies Centre, Russia, which it claims to be using for the automatic conversion of speech to text.**

### *ORGANISATIONS AND CONFERENCES*

**The Working Group on Forensic Audio (SC-03-12) of the Audio Engineering Society (AES) has recently published a second standard procedure for forensic audio. The first, AES27, was published in 1996 and provides standards for managing recorded audio materials intended for examination [46]. AES43 was published in 2000 and lays down criteria for the authentication of analogue audio tape recordings [47]. The AES Working Group is working on several additional subjects including guidelines for forensic analysis. More information can be found on its website www.aes.org. The FBI has developed its own standards for forensic audio as part of its FAVIAU (Forensic Audio, Video and Image Analysis Unit) standards.**

**Both IAFP and the ENFSI Expert Working Group for Speech and Audio Analysis organise annual conferences, frequently held back-to-back in the same venue or partly as a joint event. The proposed venue for 2002 is Russia, for 2003 Turkey. The follow-up meeting to the Martigny (1994), Avignon (1998) and Crete (2001) Speaker Recognition Tutorial and Research Workshops will be held in Toledo, Spain in 2004.**

**In December 2000, the Senior Managers of Australian and New Zealand Forensic Science Laboratories (SMANZFL) established EESAG, the Electronic Evidence Specialist Advisory Group. EESAG represents specialists involved in speech enhancement, audio and video recording analysis, image and video enhancement, and the application of digital imaging to forensic science. Its aims include the preparation of guidelines for digital image processing and for the management of recordings for the purpose of forensic examination.**

## LINGUISTIC AUHORSHIP STUDIES

**As seems to be the case for works of art in general, the authorship of literary texts is an issue that has been known to generate prolonged and sometimes downright acrimonious debate. Heated arguments have arisen over the authorship of diverse Classical Greek and Latin texts, as well as over the attribution of seventeenth-century poems and plays to authors like Shakespeare, Marlowe and Bacon [48]. In some cases of course, handwriting analysis can go a long way towards answering these questions. An example from the non-forensic domain is the study of the diary of Anne Frank [49]. However, if machine or handwriting analysis is not possible, a linguistic analysis may be the sole type of evidence that may shed light on the authorship question, short of the presence of clues in the contents of the text itself. Perhaps the first and probably still one of the few truly scientific and quantitative approaches to the study of authorship, sometimes also known as stylometry, was undertaken by Mosteller & Wallace [50] in their authorship study of *The Federalist Papers*.**

**In the last decade of the last century, a method developed by the classical scholars Morton and Michaelson called the Cusum technique enjoyed a short-lived popularity in Britain. Although the method was not officially published until 1997 [51], results of this type of analysis were readily accepted by courts in England and Australia. However, the method came in for scathing criticism in several reviews including [52, 53, 54, 55], and now seems to have vanished from the forensic scene. Also in the last decade, the importance of the availability of large databases to quantify the frequency of potentially author distinctive linguistic features came to be recognised. Two very different examples are the *KISTE*-collection of forensic texts of the BKA [56] and the *Habeas Corpus* of the University of Birmingham [57].**

**A lack of familiarity with forensic linguistic analysis combined with a tendency to rely on prejudice rather than knowledge, possibly due to an inflated sense of competence in the field of language may occasionally lead judges to formulate somewhat bizarre motivations. Coult- hard [58] reports a case where an appeal based on a careful analysis of a questioned statement caused the judges to disregard some of the challenged material, without considering it necessary to resolve the contradiction posed by their reluctance to accept either of two mutually exclusive hypotheses. This contrasts sharply with the ready acceptance of the Cusum technique by other courts, where the use of 'sophisticated' statistical methods may initially - and it now appears unjustifiably - have served to provide it with a degree of prima facie legitimacy.**

**A basic controversy that has long divided the authorship identification community is that about the distinctive character of common as opposed to less common words. While the Cusum method claimed to base its discriminatory power on differences in the frequency of very frequent phenomena like three- or four-letter words or the proportion of words starting with a vowel or a consonant, others have worked on the assumption that it is rare words that are most characteristic of a person's style. More generally, the idea is to establish author distinctive features. Most recently, attempts have been undertaken to use neural networks to detect systematic differences between authors. One major drawback for the forensic context is that meaningful results tend to presuppose a fairly large amount of language. Unfortunately, in many cases, forensic texts tend to be extremely short [59]. A useful survey of the type of information that may be relevant in linguistic authorship identification is provided by McMenamin [60]. Woolls and Coulthard [61] describe the use of a series of computer programmes specially developed to deal with forensic material in questions of disputed authorship, including suspected plagiarism. Chaski [62] describes the results of an attempt to provide empirical tests for author identification following recent court decisions in the US on the admissibility of language-based authorship identification. A thorough treatment of some of the theoretical problems in authorship identification is given in [63].**

***ORGANISATIONS***

**The International Association of Forensic Linguistics (IAFL) was founded in1991. In addition to authorship attribution, forensic linguistics includes the study of courtroom discourse, courtroom interpreting and translation, comprehensibility of legal documents and texts, including the police caution issued to suspects, and the use of linguistic evidence in court. IAFL organises annual conferences, maintains a website (www.iafl.org) and publishes the journal *Forensic Linguistics: The International Journal of Speech, Language and the Law* with its sister organisation IAFP.**

## CONCLUSIONS

### *SPEAKER IDENTIFICATION*

**Although the performance levels obtained by state-of-the-art speaker recognition technology are now comparable to those of other major biometric identification methods [64], prevailing conditions in the forensic context have so far stood in the way of large-scale introduction of automatic methods [65]. Variations in recording and transmission conditions, the presence of background noise and of variation due to differences in communicative context are responsible for performance degradations of such severity that automatic methods either cannot be applied or their results are difficult to interpret. Meanwhile, forensic speaker identification practice continues to be heavily dominated by the use of a wide variety of largely subjective procedures of which many have a strong phonetic or acoustic basis. The need to validate these methods is increasingly acknowledged within organisations like IAFP and the ENFSI Expert Working Group for Forensic Speech and Audio Analysis. Recent developments in the interpretation of the evidential value of forensic evidence are also beginning to make themselves felt in the forensic speaker identification community.**

**Guidelines have been suggested for the conduct of earwitness identification tests. They are similar in purpose and scope to those advocated for the more widely studied field of visual identification by witnesses.**

### *INTEGRITY AND AUTHENTICITY EXAMINATIONS*

**A promising development in the field of authenticity and integrity examinations of audio recordings in the analogue domain is the use of Faraday crystals as pioneered by a number of Russian scientists. This potential gain is offset by the widespread availability of relatively inexpensive digital sound processing equipment. Its ease of operation makes certain types of manipulation comparatively easy to perform. If done competently, such manipulation may leave no traces. As part of the chain-of-custody process, audio recordings, like all digital data, are therefore increasingly required to be authenticated by means of checksums and hash codes or other methods to ensure their integrity. The formulation of standards for the forensic examination of audio recordings as undertaken by the AES is a useful initiative, which may serve to improve standards across the whole field of forensic audio examination.**

### *LINGUISTIC AUTHORSHIP ATTRIBUTION*

**Authorship attribution is probably the oldest application of forensic linguistics. Nevertheless, the discriminatory power of the methods used so far remains relatively weak, if it has not been shown to be totally lacking. In countries like the United States of America and Australia, other applications of forensic linguistics, not concerned with authorship identification, are becoming increasingly prominent, as witness publications in journals like *Forensic Linguistics*.**

*FINAL CONCLUSION*

**On the basis of the findings of the survey, it would appear that the volume of work undertaken in forensic speech and audio analysis has clearly increased over the last years. There are signs that recent developments in the interpretation of the evidential value of forensic evidence are also beginning to make themselves felt in the forensic speaker identification community. More importantly, there are also clear indications of a growing awareness among those working in the field of forensic speech and audio analysis of the need to view validation of the methods used as an integral part of their discipline. In a field that was - and some would argue still is - somewhat controversial, these developments may be long overdue but that does not make them any less welcome.**

**REFERENCES:**


[1] Boves L (1998) 'Commercial Applications of Speaker Verification: Overview and Critical Success Factors', *Proceedings of RLA2C Workshop on Speaker Recognition and its Commercial and Forensic Applications*, Avignon, 150-159.

[2] Confino J (2001) '"Listen to the Customers": Implementation of a Speaker Verification System in the Bank Industry', paper presented at *2001- A Speaker Odyssey*, Crete, Greece.

[3] Hollien H, Bennett G & Gelfer MP (1983) 'Criminal Investigation Comparison: Aural versus Visual Identification Resulting from a Simulated Crime', *Journal of Forensic Sciences* 28, 208-222.

[4] McGehee F (1937) 'The Reliability of the Identification of the Human Voice', *Journal of General Psychology*, 17: 249-271.

[5] Thompson C (1985) 'Voice Identification: Speaker Identifiability and a Correction of the Record regarding Sex Effects', *Human Learning* 4, 19-27.

[6] Broeders APA (1996) 'Earwitness Identification: Common Ground, Disputed Territory and Uncharted Areas', *Forensic Linguistics* 3(1), 3-13.

[7] Broeders APA & Van Amelsvoort AG (1999) 'Line-up Construction for Forensic Earwitness Identification: a Practical Approach', in: *Proceedings of the XIVth International Congress of Phonetic Sciences,* San Francisco, 1373-1376.

[8] Nolan F & Grabe E (1996), ‘Preparing a Voice Line-up’, *Forensic Linguistics* 3(1), 74-94.

[9] Stuart Laubstein A (1997), ‘Problems of Voice Line-ups’, *Forensic Linguistics* 4(2), 262-279.

[10] Van Amelsvoort AG (1999) *Handleiding Confrontatie*, Elsevier bedrijfsinformatie: Den Haag.

[11] Wells GL, Small M, Penrod S, Malpass RS, Fulero SM, & Brimacombe CAE (1998), 'Eyewitness Identification Procedures: Recommendations for Line-ups and Photo- spreads', *Law and Human Behavior* 22 (6), 603-647.

[12] Yarmey AD (1995) 'Earwitness and Evidence Obtained by Other Senses' in: Bull R & Carson D (eds.) *Handbook of Psychology in Legal Contexts*, Wiley: Chichester, 261-273.

[13] Yarmey AD, Yarmey AL, Yarmey MJ & Parliament L (2001) 'Common Sense Beliefs and the Identification of Familiar Voices', *Applied Cognitive Psychology* 15(3), 283-299.

[14] Hollien H (1996) 'Consideration of Guidelines for Earwitness Line-ups', *Forensic Linguistics* 3 (1), 14-23.

[15] Bull R & Clifford BR (1999) 'Earwitness Testimony', *Medicine, Science and Law* 39 (2), 120-127.

[16] Collins B & Mees I (1998) *The Real Professor Higgins: The Life and Career of Daniel Jones*, Mouton de Gruyter: Berlin.

[17] Bolt RH et al., (1979) *On the Theory and Practice of Voice Identification*, National Academy of Sciences: Washington DC.

[18] Saks MJ (1998) 'Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science', Hastings Law Journal 49(4), 1069-1141.

[19] Braun A & Künzel HJ (1998) 'Is Forensic Speaker Identification Unethical - or Can it be Unethical *not* to Do it?', *Forensic Linguistics* 5(1), 10-21.

[20] Meuwly D (2001) *Reconnaissance de Locuteurs: l'Apport d'une Approche Automatique*, PhD Thesis, University of Lausanne.

[21] Schmidt-Nielsen A & Crystal TH (1998) 'Human vs. Machine Speaker Identification with Telephone Speech', *Proceedings ICSLP* '98.

[22] Andrews WA, Kohler MA & Campbell JP (2001) 'Phonetic, Idiolectal and Acoustic Speaker Recognition', paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.

[23] Doddington G (2001) 'Speaker Recognition Based on Idiolectal Differences between Speakers', to be published in *Eurospeech*, September 2001.

[24] Marescal F (1999) 'The Forensic Speaker Recognition Method Used by the French Gendarmerie', internal publication, IRCGN: Paris.

[25] Pfister B (2001) 'Personenidentifikation anhand der Stimme', *Kriminalistik* 55(4), 287-292.

[26] González-Rodriguez J, Ortega-García J & Lucena-Molina J (2001) 'On the Application of the Bayesian Framework to Real Forensic Conditions with GMM-based Systems', paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.

[27] Nakasone H & Beck SD (2001) 'Forensic Automatic Speaker Identification', paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.

[28] Evett IW (1991) 'Interpretation: A Personal Odyssey', in: Aitken CGG & DA Stoney, *The Use of Statistics in Forensic Science*, Ellis Horwood: New York, 9-22.

[29] Evett IW (1998) 'Toward a Uniform Framework for Reporting Opinions on Forensic Science Casework', *Science and Justice* 38(3), 198-202.

[30] Evett IW (1995) 'Avoiding the Transposed Conditional', *Science & Justice* 35(2), 127-131.

[31] Champod C & Meuwly D (2000) 'The Inference of Identity in Forensic Speaker Recognition', *Speech Communication* 31, 193-203.

[32] Aitken CGG & Taroni F (1998) 'A Verbal Scale for the Interpretation of Evidence', *Science and Justice* 38(4), 279-281.

[33] Meuwly D & Drygajlo A (2001) 'Forensic Speaker Recognition Based on a Bayesian Framework and Gaussian Mixture Modelling', paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.
[34] Boves L & Koolwaaij J (1999) 'On Decision Making in Forensic Casework', *Forensic Linguistics* 6(2), 242-264.
[35] Broeders APA (1999) 'Some Observations on the Use of Probability Scales in Forensic Identification', *Forensic Linguistics* 6(2), 228-241.
[36] Found B & Rogers D (2000) 'The Development of a Program for Characterising and Profiling Individual and Collective Expertise in Forensic Handwriting Examination', *Abstracts EAFS 2000*, Cracow, 136.
[37] GFCP (1999) 'Petition pour l'Arrêt des Expertises Vocales (sans Validations Scientifiques)', www.lapetition.com.
[38] Boé L-J (2000) 'Forensic Voice Identification in France', *Speech Communication* 31, 205-224.
[39] GFCP (1991) 'About the Ethics of Speaker Identification', *Proceedings of the XIIth International Congress of Phonetic Sciences*, Aix-en-Provence, Vol. 1, 397.
[40] Broeders APA (1991) 'Great Debate on….', *Nesca – The ESCA Newsletter*, No. 4, 50-51.
[41] Wright P (1988) *Spycatcher: the Candid Autobiography of a Senior Intelligence Officer*, Boston, Mass.: GK Hall.
[42] Boss D, Gfroerer S, Neoustroev N (2001) 'A New Tool for the Visualisation of Magnetic Features on Tapes', paper presented at the IAFP-ENFSI.EWG Joint Meeting, Rosny-sous-Bois, France.
[43] Koenig BE (1990) 'Authentication of Forensic Audio Recordings', *Journal of the Audio Engineering Society* 38, 1/2, 3-33.
[44] Dean DJ (1991) 'The Relevance of Replay Transients in the Forensic Examination of Analogue Magnetic Tape Recorders', internal publication, Home Office, PSDB, St. Albans.
[45] French JP (1990) 'Analytic Procedures for the Determination of Disputed Utterances', in: Kniffka H (ed.) *Texte zu Theorie und Praxis forensischer Linguistik*, Niemeyer, Tübingen, 201-213.
[46] AES (1996) 'AES27-1996: AES Recommended Practice for Forensic Purposes - Managing Recorded Audio Materials Intended for Examination' *Journal of the Audio Engineering Society* 44(4), 274-283.
[47] AES (2000) 'AES43-2000: AES Standard for Forensic Purposes - Criteria for the Authentication of Analog Audio Tape Recordings' *Journal of the Audio Engineering Society* 48(3), 204-214.
[48] Holmes DI (1998) 'The Evolution of Stylometry in Humanities Scholarship' *Literary and Linguistic Computing* 13(3), 111-117.
[49] Hardy HJJ (1989) 'Document Examination and Handwriting Identification of the Text Known as the Diary of Anne Frank: Summary of Findings', in: Barnouw D and Van der Stroom G (eds.) *The Diary of Anne Frank: the Critical Edition*, New York: Bantam, Doubleday, Dell Publishing Group, 102-165.

**[50] Mosteller F & Wallace DL (1984), *Applied Bayesian and Classical Inference in the Case of the Federalist Papers*, 2nd edition, New York: Springer Verlag.**

**[51] Farringdon JM (1996) *Analysing for Authorship: A Guide to the Cusum Technique*, Cardiff: University of Wales Press.**

**[52] Canter D (1992) 'An Evaluation of the "Cusum" Stylistic Analysis of Confessions', *Expert Evidence* 1(2), 93-99.**

**[53] De Haan P & Schils E (1994), 'The Cusum Plot Exposed', in: Frier U et al. (eds.), *Creating and Using English Language Corpora*, Rodopi: Amsterdam, 93-105.**

**[54] Hardcastle RA (1997), 'CUSUM: a credible method for the determination of authorship?', *Science & Justice* 37(2), 129-138.**

**[55] Barr GK (1998) 'The Cusum Mechanism – A Review of Analyzing for Authorship by Jill M Farringdon', *Expert Evidence* 6, 43-55.**

**[56] Schall S & Hehn W (1996) 'Das System KISTE im BKA', in: Kniffka H (ed.) *Recent Developments in Forensic Linguistics*, Frankfurt.**

**[57] Coulthard M (1994) 'On the Use of Corpora in the Study of Forensic Texts', *Forensic Linguistics*, 1(1), 27-43.**

**[58] Coulthard M (1997) 'A Failed Appeal', *Forensic Linguistics* 4(2), 287-302.**

**[59] Baldauf C (1999), 'Zur Signifikanz sprachlicher Merkmale im Rahmen des Autor-schaftsnachweises: Ansätze und Desiderata der forensischen Linguistik', *Archiv für Kriminologie*, 204 (3/4), 93-105.**

**[60] McMenamin GR (1993) 'Forensic Linguistics', *Forensic Science International 58*, Special issue.**

**[61] Woolls D & Coulthard M (1998), 'Tools for the Trade', *Forensic Linguistics* 5(1), 33-57.**

**[62] Chaski CE (2001) 'Empirical Evaluations of Language-Based Author Identification Techniques', *Forensic Linguistics* 8(1), 1-65.**

**[63] Grant T & Baker K (2001) 'Identifying reliable, Valid Markers of Authorship: A Response to Chaski', *Forensic Linguistics* 8(1), 66-79.**

**[64] Wayman JL (2001) 'Theory, Characterization and Testing of General Biometric Technologies' paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.**

**[65] Broeders APA (2000) 'Speech as a Biometric: Separating the Goats from the Sheep?' in: Breur CM, Kommer MM, Nijboer JF & Reijntjes JM (eds.) *New Trends in Criminal Investigation and Evidence II*, Intersentia: Antwerp, 83-94.**

**THE SURVEY**

***SCOPE OF THE SURVEY***

**In the first quarter of 2001 a questionnaire was sent to Interpol member countries and government forensic laboratories which were thought to be (considering becoming) active in the field of forensic speech and audio analysis and forensic linguistics. In all, 30 completed questionnaires were received from 21 countries. This represents a considerable increase over the results of the 1998 and 1995 surveys, when 10 and 5 replies respectively were received. Five of the 28 laboratories indicated that they were not (yet) active in the area. The responding countries are listed below, with the number in parentheses indicating the number of replies per country if greater than one:**

**Austria, Belarus, Belgium, Brazil, Canada, the Czech Republic, Finland, France (2), Germany (3), Israel, Italy, Lithuania, the Netherlands (2), Norway, Slovakia, Slovenia, Spain (3), Sweden, Switzerland (3), the United Kingdom (2), the United States.**

***CASEWORK***

**From the questionnaires received it appears that of the 30 responding laboratories by far the largest casework volume in *speaker identification* is reported by the Lithuanian Institute of Forensic Investigation, which carries out many hundreds of speaker identification tests per year. The Institute of Criminalistics in Prague (the Czech Republic), the Policía Científica in Madrid (Spain), the State Expert and Forensic Science Centre of the Ministry of the Internal Affairs (Republic of Belarus) and the FBI (USA) report slightly over a hundred cases per year, with the FBI practising speaker identification for investigative purposes only. The German BKA reports approximately the same number. The NFI (the Netherlands) and the Belgian Federal Police report between 50 and 80 cases per year. Laboratories using automatic methods typically perform relatively few cases, with IRCGN (Gendarmerie Nationale, France) and ETH Zürich (Switzerland) both reporting ten cases, and ISPC-UNIL in Lausanne (Switzerland) and the Guardia Civil (Spain) both reporting just five. Neither of the two reporting laboratories in the UK, the Forensic Science Service and the Metropolitan Police Forensic Audio Laboratory, are active in the area of speaker identification, the latter laboratory reporting that this type of work is carried out by UK based (private) experts, independently of the Metropolitan Police.**

**By far the largest volume of *audio enhancement* work is reported by the Metropolitan Police Forensic Audio Laboratory, who process close to 3000 recordings on an annual basis, with the FBI Forensic Audio Laboratory following at a distance with some 600 cases a year. However, the figures reported may be slightly misleading in that in many countries like the Netherlands the bulk of the audio (and video) enhancement work is carried out by regional police laboratories rather than by the specialist forensic laboratories approached for this review.**

**Both speaker identification by earwitnesses and linguistic authorship identification are relatively rarely reported, with the BKA as the exception. It reports about 150 cases per year in the latter category.**

***RESEARCH***

**Research projects in the area of speaker identification are reported by the following institutes:**
**BKA (Germany): several projects, including application of Faraday and Kerr type effects to authentication work and work on GSM-type telephone speech;**
**ETH Zürich (Switzerland): ongoing research on automatic speaker recognition;**
**FBI (USA): PC-based *FASR* (Forensic Automatic Speaker Recognition) and *MMI* (Magnetic Media Analyser);**
**Gendarmerie Criminal Department (Turkey): ongoing research on the *KASIS* software;**
**Guardia Civil (Spain): three-year project with Polytechnic University of Madrid on *IdentiVox* system;**
**Institute of Acoustics of the Austrian Academy of Sciences: STX software system;**
**Institute of Forensic Examination (Lithuania): ongoing research on speaker identification;**
**Israel National Police: work on GMM-based speaker recognition;**
**IPSC Lausanne (Switzerland): user-friendly interfaces for software and recognition methods focused on GSM-signals, collaboration with EPFL-DE-LTS of the Swiss Federal Institute of Technology;**
**IRCGN (France): ongoing research on automatic speaker recognition;**
**Metropolitan Audio Laboratory (UK): ongoing research on analogue and digital audio authenticity;**
**NBI (Finland): several projects with the University of Helsinki;**
**Netherlands Forensic Institute (the Netherlands) reports work on earwitness identification procedures, audio authenticity examinations and audio casework examination protocols;**
**Policía Cientifíca (Spain): ongoing work on specific speech features;**
**RaCIS Carabinieri (Italy): ongoing research with Fondazione Ugo Bordoni;**

*DATABASES*

The following databases were reported:
BKA (Germany): *DRUGS* (Databank Regionale Umgangssprache), *KISTE* (linguistic authorship identification) and *TELDAT* (telephone signal parameters);
ETH Zürich (Switzerland): speech signal database;
FBI (USA): collection of gunshot analysis recordings;
Guardia Civil (Spain): *Ahumada/Gaudi* corpus of some 450 speakers;
Institute of Acoustics of the Austrian Academy of Sciences (Austria): corpus of different languages including Albanian, Bosnian, German, Igbo, Romanian;
Institute of Criminalistics Prague (Czech Republic): anonymous calls;
IPSC Lausanne (Switzerland): corpus of 16 pairs of French speaking soundalikes;
IRCGN (France): database of 250 male and 150 female speakers of French;
Israel National Police: Hebrew speakers, magnetic stop/start events;
Netherlands Forensic Institute (the Netherlands): corpus of spoken Dutch (*CGN*);
Policía Científíca (Spain): speaker database *LOCOPOL*;
RaCIS Carabinieri (Italy): *IDEM* formants database;

*EDUCATION AND TRAINING*

The FBI (USA) reports progress in digital evidence handling procedures. Most laboratories provide forms of in-house and on-the-job training.

*QUALITY ASSURANCE*

As in forensic science in general, the introduction of quality assurance procedures is an issue that is becoming more and more important. In Europe, this work is taken forward within ENFSI, the European Network of Forensic Science Institutes. This organisation was formally established in 1994 and seeks to promote education and training of experts, the introduction and enforcement of quality assurance systems and the harmonisation of methods and techniques in the various forensic disciplines. Member laboratories like the FSS (the Forensic Science Service) in Britain, SKL in Sweden, NBICL in Finland and NFI in the Netherlands have certified many of their forensic examinations with nationally operating, external and independent laboratory certification boards, such as UKAS in the United Kingdom and the Council for Accreditation in the Netherlands, and are continuing to do so. An increasingly important role in this context is being played by the ENFSI Expert Working Groups set up in the last decade. Like their American counterparts, such as the Scientific Working Group for Materials Analysis (SWGMAT) and the Scientific Working Group for Document Examination (SWGDOC) working under the auspices of the FBI, and similar groups in Australia and New Zealand, such as the Scientific Advisory Groups (SAGs) operating within the context of SMANZFL (the Senior Managers of Australian and New Zealand Forensic Science Laboratories), many of the ENFSI Expert Working Groups, such as the Drugs, Fibres, Paint, Firearms and DNA Groups are actively involved in drawing up best practice manuals, setting up collaborative tests and education and training

**programmes and working towards increased harmonisation and standardisation of methods and techniques. Unfortunately, given the wide variety of procedures and practices in forensic speaker identification to which the present survey also bears witness, harmonisation and quality assurance will not be easy to achieve in this area within the near future.**

**IRCGN (France) is the only laboratory to report that it is preparing its speaker identification method for accreditation. IAFP has put in place an accreditation procedure for practising forensic phoneticians. So far two individuals have successfully completed this procedure. A number of laboratories report that the institute as a whole is seeking to comply with ISO 17025 for accreditation (National Police Laboratory - Israel, and Guardia Civil - Spain). The Metropolitan Police Audio Laboratory is registered according to ISO 9002.The BKA reports that it is planning proficiency tests for all German state and federal labs for the year 2001. Many other laboratories report work on SOP's (standard operating procedures) and examination protocols.**

**Within the field of forensic audio, harmonisation and standardisation are probably much easier to achieve than for forensic speaker identification. The AES Standards for forensic audio provide a clear indication of this. In Australia and New Zealand, the newly established EESAG is also committed to playing a key role in promoting and developing mechanisms of quality management and training.**

***PUBLICATIONS***

**The following is a list of publications compiled from the questionnaires and complemented with some additional articles published during the survey period:**

***Forensic Speech and Audio Analysis - Publications***

**AES (1996) 'AES27-1996: Recommended Practice for Forensic Purposes - Managing Recorded Audio Materials Intended for Examination' *Journal of the Audio Engineering Society* 44(4), 274-283.**

**AES (2000) 'AES43-2000: AES Standard for Forensic Purposes - Criteria for the Authentication of Analog Audio Tape Recordings' *Journal of the Audio Engineering Society* 48(3), 204-214.**

**Bobda AS, Wolf H-G, Peter L (1999), 'Identifying Regional and National Origin of English-speaking Africans Seeking Asylum in Germany', *Forensic Linguistics* 6(2), 300-319.**

**Boé L-J (2000) 'Forensic Voice Identification in France', *Speech Communication*, 31, 205-224.**

**Bouten JS & Broeders APA (1999) 'Text-Independent Forensic Speaker Identification Using Telephone Speech', *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 1377-1379.**

**Braun A (1998) 'Die forensische Analyse von Stimme und Sprache', in: Gundermann H (ed.) *Die Ausdruckswelt der Stimme. Erste Stuttgarter Stimmtage*, Heidelberg: Hüthig, 88-102.**

**Braun A & Cerrato L (1999) 'Estimating Speaker Age Across Languages', *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 1369-1372.**

**Braun A & Köster J-P (2000) 'Speaker Identification by Untrained and Trained Listeners', in: *Proceedings of the American Academy of Forensic Sciences Annual Meeting*, Reno, 119.**

**Braun A & Köster J-P (2000) 'Auditive Sprechererkennung durch geübte und ungeübte Hörer', in: Hoffmann R (ed.): *Phonetik - Sprachkommunikation - Rehabilitations-technik. Professor Dieter Mehnert zum 65. Geburtstag*. Dresden: web Universitäts-verlag, 41-52.**

**Braun A & Künzel H (1998) 'The Influence of Alcohol on Speech Production: Phonetic and Linguistic Aspects, *Proceedings of the American Academy of Forensic Sciences 50th Anniversary Meeting*, San Francisco, 110**
**.**
**Broeders APA (1999) 'Foreword', *Forensic Linguistics* 6(2), 211-213.**

**Broeders APA (1999) 'Some Observations on the Use of Probability Scales in Forensic Identification', *Forensic Linguistics* 6(2), 228-241.**

**Broeders APA (2000) 'Speech as a Biometric: Separating the Goats from the Sheep?' in: Breur CM, Kommer MM, Nijboer JF & Reijntjes JM (eds.) *New Trends in Criminal Investigation and Evidence II*, Intersentia: Antwerp, 83-94.**

**Broeders APA (2000) 'Forensic Speech and Audio Analysis: the State of the Art in 2000 AD', *Actas del I Congreso de la Sociedad Española de Acústica Forense*, Madrid, 13-24.**

**Broeders APA & Van Amelsvoort AG (1999) 'Lineup Construction for Forensic Earwitness Identification: a Practical Approach', *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 1373-1376.**

**Broeders APA & Van Amelsvoort AG (2001) ' A Practical Approach to Forensic Earwitness Identification: Constructing a Voice Lineup', *Proceedings of the 2nd European Academy of Forensic Science Meeting*, Cracow.**

**Campbell JP (1997) 'Speaker Recognition: A Tutorial', *Proceedings of the IEEE* 85(9), 1437-1462.**

**Champod C & Meuwly D (2000) 'The Inference of Identity in Forensic Speaker Recognition', *Speech Communication*, 31, 193-203.**

**DeJong G (1998) *Earwitness Characteristics and Speaker Identification Accuracy*, unpublished PhD Dissertation, University of Florida.**

**Delgado-Romero C (2001) *La Identificación de Locutores en el ámbito forense,* PhD thesis, Universidad Complutense de Madrid.**

**Eskelinen-Rönkä P & Niemi-Laitinen T (1999) 'Testing Voice Quality Parameters in Speaker Recognition', *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 149-151.**

**Foulkes P & Barron A (2000), 'Telephone Speaker Recognition Amongst members of a Close Social Network', *Forensic Linguistics* 7(2), 180-198.**

**French JP (1998) 'Mr Akbar's Nearest Ear versus the Lombard Reflex: a Case Study in Forensic Phonetics', *Forensic Linguistics* 5(1), 58-68.**

**Gfroerer S (1998) 'Kriminalwissenschaftliche Möglichkeiten der Sprecheranalyse', in: *Entführung - Kriminalistische Aspekte der Ermittlungsführung und operativer Maßnahmen*. Forschungsbericht. Polizeiführungsakademie Münster-Hiltrup, 170-189.**

**Gfroerer S (2000) Report of the 3rd Meeting of the ENFSI Working Group for Forensic Speech and Audio Analysis, Cracow, 12.-16.09, *The Phonetician*, 82, 43-46.**

**Gfroerer S & Baldauf C (2000) 'Sprechererkennung, Tonträgerauswertung und Autorenerkennung', in: Beleke N (ed.) *Kriminalistische Kompetenz. Kriminalwissenschaften, kommentiertes Recht und Kriminaltaktik für Studium und Praxis*, Kap. II, 14: Kriminaltechnik, 3-16.**

**González-Rodriguez J, Ortega-García J & Lucena-Molina J (2001) 'On the Application of the Bayesian Framework to Real Forensic Conditions with GMM-based Systems', paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.**

**Greisbach R (1999) 'Estimation of Speaker Height from Formant Frequencies', *Forensic Linguistics* 6(2), 265-277.**

**Güner L & Malkoç E (1998) 'Forensic Applications in Speaker and Voice Recognition' *Proceedings of the 8th COST Workshop*, Ankara.**

**Hollien H & Schwartz R (2000) 'Aural-perceptual Speaker Identification: Problems with Non-contem- porary Samples', *Forensic Linguistics* 7(2), 199-211.**

**Hollien H & Schwartz R (2001) 'Speaker Identification Utilising Non-contemporary Speech', *Journal of Forensic Sciences* 46(1), 63-67.**

**Koenig BE, Hoffman SM, Nakasone H & Beck SD (1998) 'Signal Convolution of Recorded Free-Field Gunshot Sounds', *Journal of the Audio Engineering Society* 46(7/8), 634-653.**

**Koolwaaij J (2000) *Automatic Speaker Verification in Telephony: A Probabilistic Approach*, PhD Thesis, University of Nijmegen.**

**Köster O (2001) 'Die Datenbank regionaler Umgangssprachen (DRUGS). Ein neues Datenbank-Expertensystem für die forensische Sprechererkennung' *Kriminalistik* 55 1/01, 46-50.**

**Köster O, Hess M, Schiller O and Künzel H (1998) 'The Correlation Between Auditory Speech Sensitivity and Speaker Recogniton Ability, *Forensic Linguistics* 5(2), 22-32.**

**Kredens K & Góralewski-Ğach G (1998) 'Language as Sole Incriminating Evidence: the Agustynek Case', *Forensic Linguistics*, 5(2), 193-202.**

**Künzel HJ (2000) 'Effects of Voice Disguise on Speaking Fundamental Frequency', *Forensic Linguistics* 7(2), 149-179.**

**Künzel HJ (2001) 'Beware of the "Telephone Effect": the Influence of Telephone Transmission on the Measurement of Formant Frequencies', *Forensic Linguistics* 8(1), 80-99.**

**Künzel HJ (2001) 'Eine Datenbank regionaler Umgangssprachen des Deutschen (DRUGS) für forensische Anwendungen', in: *Zeitschrift für Dialektologie und Linguistik ZDL* 2, 129-154.**

**Lipeika A & Lipeikienā J (1999) 'Speaker Recognition Based on the Use of Vocal Tract and Residue Signal LPC Parameters, *INFORMATICA*, 10(4), 449-456.**
**Lucena-Molina J (2000) 'Técnicas Propuestas para Autenticación de Grabaciones en Soportes Magné-ticos y su Aplicación Forense' *Actas del I Congreso de la Sociedad Española de Acústica Forense*, Madrid, 43-53.**

**Lucena-Molina J & Díaz-Gómez JJ (2000) 'Evaluación del Sistema de Reconocimiento Automático de Locutores *IdentiVox 2000* con la Base de Datos AHUMADA, *Actas del I Congreso de la Sociedad Española de Acústica Forense*, Madrid, 183-192.**

**Markham D (1999) 'Listeners and Disguised Voices: The Imitation and Perception of Dialectal Accent', *Forensic Linguistics* 6(2), 289-299.**

**Meuwly D (2000) 'Voice Analysis', *Encyclopedia of Forensic Sciences*, Academic Press. Meuwly D (2001) *Reconnaissance de Locuteurs: l'Apport d'une Approche Automatique*, PhD Thesis, University of Lausanne.**

**Meuwly D & Drygajlo A (2000) 'The Influence of the Telephone Network on Automatic Forensic Speaker Recognition', *Proceedings of the 2nd European Academy of Forensic Science Meeting*, Cracow.**

**Meuwly D, El Maliki M, & Drygajlo A (1998) 'Forensic Speaker Recognition Using Gausian Mixture Models and a Bayesian Framework', *Proceedings of the 8th COST Workshop*, Ankara, 52-55.**

**Moosmüller S (2001) 'The Influence of Creaky Voice on Formant Frequency Changes', *Forensic Linguistics* 8(1), 100-112.**

**Niemi-Laitinen T, Iivonen A & Harinen K (1999) 'Similarity Degree between Speakers on the Basis of Short FFT Spectra', *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 153-156.**

**Ortega-García J, González-Rodríguez J & Marrero-Aguilar V (2000) 'AHUMADA: a Large Speech Corpus in Spanish for Speaker Characterisation and Identification', *Speech Communication*, 31(3), 255-264.**

**Pfister B (2001) 'Personenidentifikation anhand der Stimme', *Kriminalistik* 55(4), 287-292. Schiller N and Köster O (1998) 'The Ability of Expert Witnesses to Identify Voices: a Comparison between Trained and Untrained Listeners', *Forensic Linguistics* 5(1), 1-9.**

**Solewicz YA (2001) 'Noise Robustness in Forensic Speaker Verification', paper presented at *2001 - A Speaker Odyssey*, Crete, Greece.**

**Thomas DB (2001) 'Echo Correlation Analysis and the Acoustic Evidence in the Kennedy Assassination Revisited', *Science & Justice*, 41(1), 21-32.**

**Wagner I and Köster O (1999) 'Perceptual Recognition of Familiar Voices using Falsetto as a Type of Voice Disguise', *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 1381-1385.**
**Yarmey AD (2001) 'Earwitness Descriptions and Speaker Identification', *Forensic Linguistics* 8(1), 113-122.**

*Forensic Speech and Audio Analysis - Papers*

**Delgado-Romero C (1999) 'Tecnicas digitales de análisis audiovisual en Acústica Forense', paper presented at the IIIrd Congreso de Investigadores Audiovisuales "Los Medios del Tercer Milenio", Universidad Complutense de Madrid.**

**Gfroerer S (1998) 'Forensic Speaker Recognition - An Overview', paper presented at the THESSEUS conference. Athens, Greece.**

**Gfroerer S (1999) 'Forensische Sprachverarbeitung und Sprachverbesserung bei der Elektronischen Aufklärung', paper presented at the 2nd Seminar "Electronic Surveillance". Tomar, Portugal.**

**Gfroerer S (1999) 'Intelligibility Enhancement and the Transcription of Disputed Utterances – A Survey', paper presented at the 2nd Meeting of the ENFSI Working Group for Forensic Speech and Audio Analysis, Madrid, Spain.**

**Stenberg M (2000) 'Desideria: a Case Report on Identification by Voice Line-up', paper presented at the 2nd European Academy of Forensic Science Meeting, Cracow.**

*Forensic Linguistics - Publications*

**Baldauf C (1999) 'Sprachliche Spurensuche bei der Aufklärung von Erpressungsstraftaten', *Polizei Heute* 3, 86-90.**

**Baldauf C (1999) 'Zur Signifikanz sprachlicher Merkmale im Rahmen des Autorschaftsnachweises: Ansätze und Desiderate der forensischen Linguistik', *Archiv für Kriminologie*, 204/3,4, 93-105.**

**Baldauf C (ed.) (2000) *Autorenerkennung*. Symposium des Bundeskriminalamtes, 03.-05. April 2000. Wiesbaden.**

**Baldauf C & Stein S (2000) 'Die Bullen sollen nicht auf blöde Gedanken kommen - Phraseologismen in Erpresserbriefen', *Kriminalistik* 10/00, 666-670.**

**Gfroerer S & Baldauf C (2000) 'Sprechererkennung, Tonträgerauswertung und Autorenerkennung', in: Beleke N (ed.) *Kriminalistische Kompetenz. Kriminalwissenschaften, kommentiertes Recht und Kriminaltaktik für Studium und Praxis*. Kap. II, 14: Kriminaltechnik, 3-16.**

**Hänlein H (1999) *Studies in Authorship Recognition: A Corpus-Based Approach*, Frankfurt: Peter Lang.**

**Schall S (2000) 'Geschriebene und gesprochene Sprache bei Erpressungen, in: Baldauf, C (ed.) *Autorenerkennung*. Symposium des Bundeskriminalamtes, 03. - 05. April 2000. Wiesbaden.**

**Stein S & Baldauf C (2000) 'Feste sprachliche Einheiten in Erpresserbriefen. Empirische Analysen und Überlegungen zu ihrer Relevanz für die forensische Text-analyse', *Zeitschrift für germa- nistische Linguistik* 28/3, 377-403.**

***Forensic Linguistics - Papers***

**Baldauf C (1999) 'Forensic Linguistics in Germany', paper presented at the 4th Biennial IAFL Conference in Birmingham, 28th June – 1st July 1999.**

**Baldauf C (2000) 'Forensische Linguistik/Autorenerkennung', GAL-Tagung, 30. September 2000.**